VALUE SENSITIVE DESIGN: MACHINE LEARNING ETHICS

Kevin Mills

The Ethics Institute &

Khoury College of Computer Sciences

Northeastern University

OUR AIMS TODAY

- Technology is transforming our lives, both as individuals and as a society.
  - Many of these transformations are highly desirable (e.g. medicine; Wikipedia)
- But many aren’t, and we often recognize the problems associated with some technology only long after deploying it.
  - For example: we are currently experiencing a misinformation crisis, in large part driven by information technologies, and it’s not clear how to solve this.
    - Twitter labeled a lot of content as “misleading” during the recent election; this arguably helped combat misinformation, but raises worries about censorship.
    - What should Twitter have done? They were essentially forced to address an extremely difficult ethical question because they control an important piece of technology that inevitably intersects with these ethical issues.
    - Note that for Twitter, doing nothing is still a decision! It is the decision to allow people to use your platform in certain ways.
      - (Often a so-called “neutral” decision is a disguised ethical decision. It might still be the right decision, but that’s an ethical question.)

VALUES AND TECHNOLOGY

Technology is the result of human imagination

All technology involves design

All design involves choices among possible options

All choices reflects values

Therefore, all technologies reflect and affect human values

Ignoring values in the design process is irresponsible

Making ethical judgments is sometimes an unavoidable part of producing technology, and the stakes can be high!

OUR AIMS TODAY

- Our aim is to equip future producers of technology (that's you) with the tools they need to confront the ethical challenges they are likely to face in their careers.

- Our aim is **not** to tell you what to think or do.

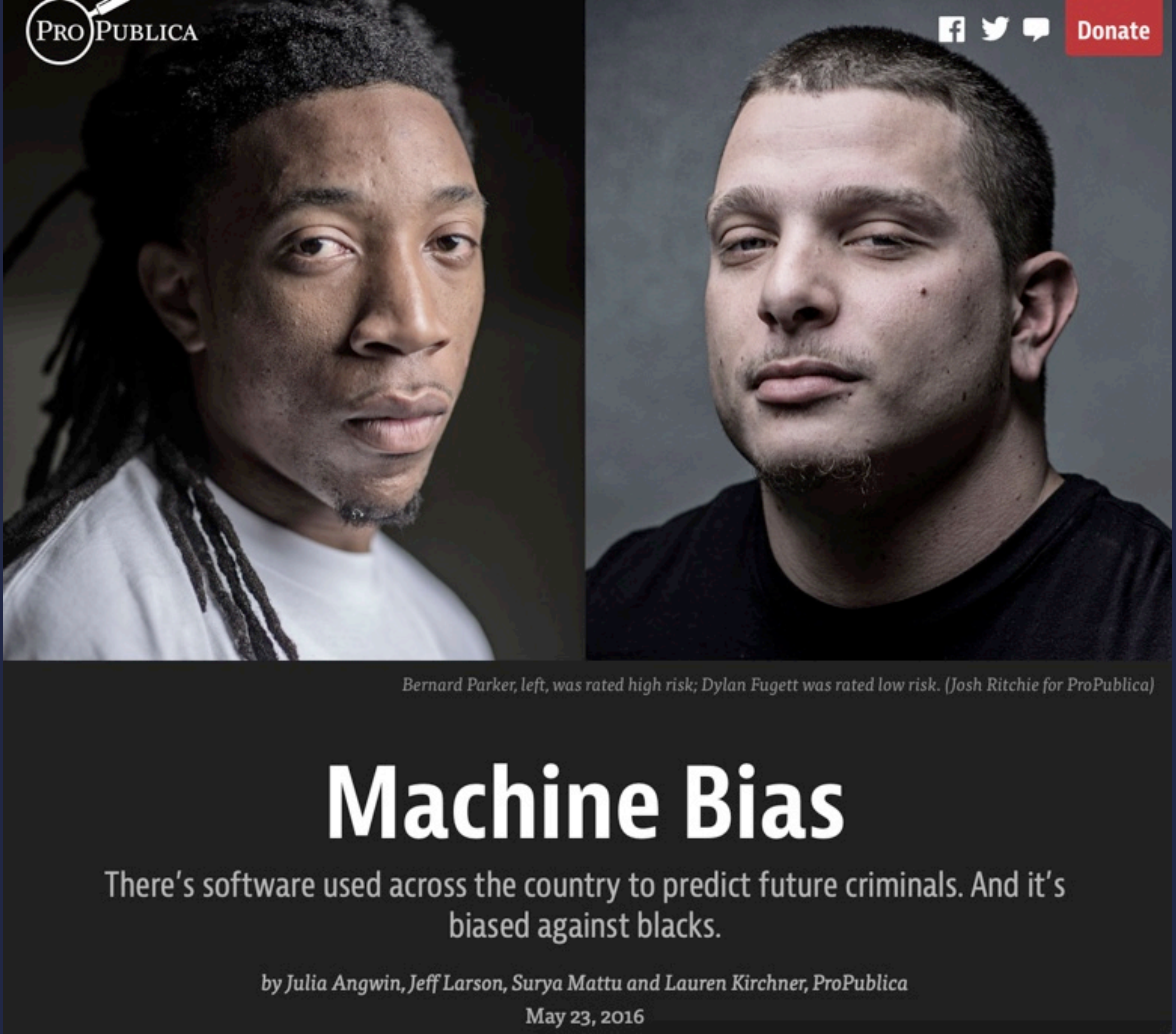
    - Nothing I say today should be taken as gospel.

- Our aim is to help you think through ethical dilemmas for yourselves.

    - Although, it bears emphasizing that when you are presented with a particularly tricky ethical dilemma, it is often worth reaching out to people with relevant expertise.

    - For example, I don't know what kinds of effects labeling some content on Twitter as "misleading" has (maybe certain demographics will take content labeled as "misleading" to be more likely to be true), and e.g. psychologists and anthropologists might be able to help me with this question.

MACHINE LEARNING ETHICS: SOME CASE STUDIES



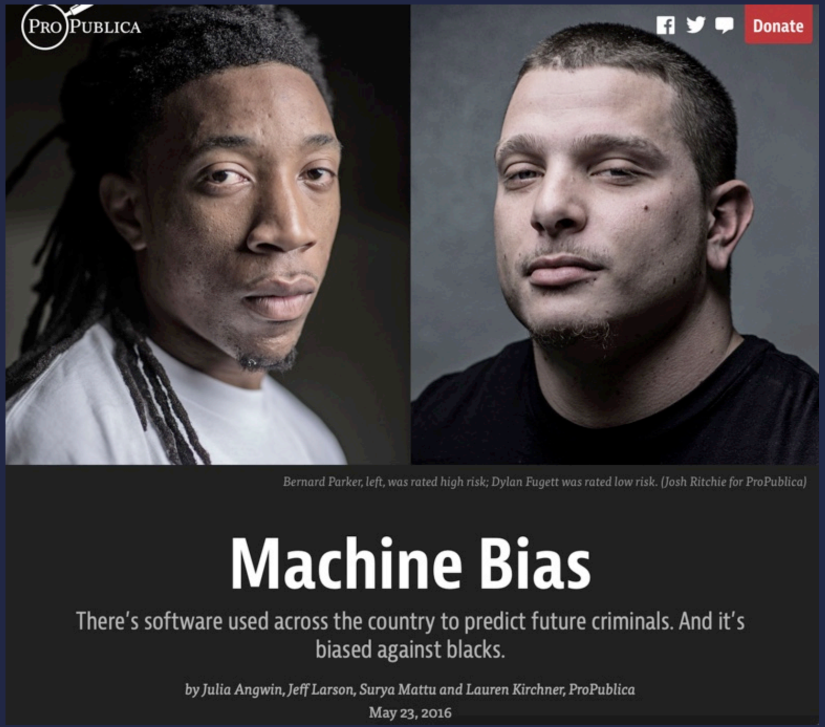
Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016



PROPUBLICA

Donate

Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

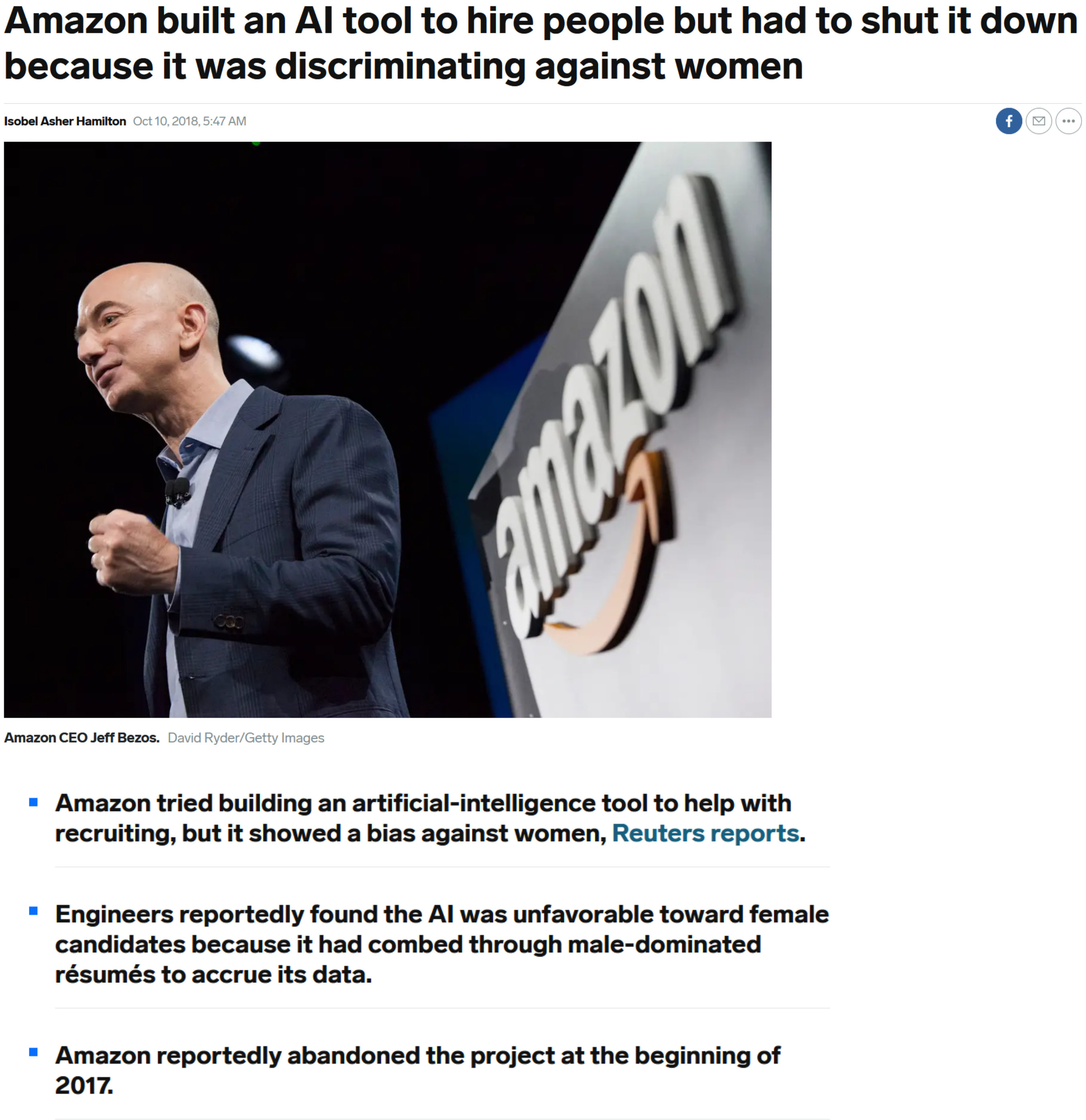
May 23, 2016

	White	Black
Labeled Higher Risk & Didn't Re-Offend (False +s)	23. %	44.9%
Labeled Lower Risk & Did Re-Offend (False -s)	47.7%	28.0%

Unfair distribution of error by racial class membership

ProPublica examined COMPAS risk assessment scores for 7000 defendants in 2013-2014 in Broward County, FL. Propublica found that while COMPAS correctly predicted recidivism 61% of the time, **the likelihood of different types of errors (FP and FNs) differed by race** (see table above)

Northpointe responded that COMPAS was nevertheless fair because its positive predictions of recidivism were correct at the same rates regardless of racial group membership; this response was followed by a detailed rebuttal by ProPublica



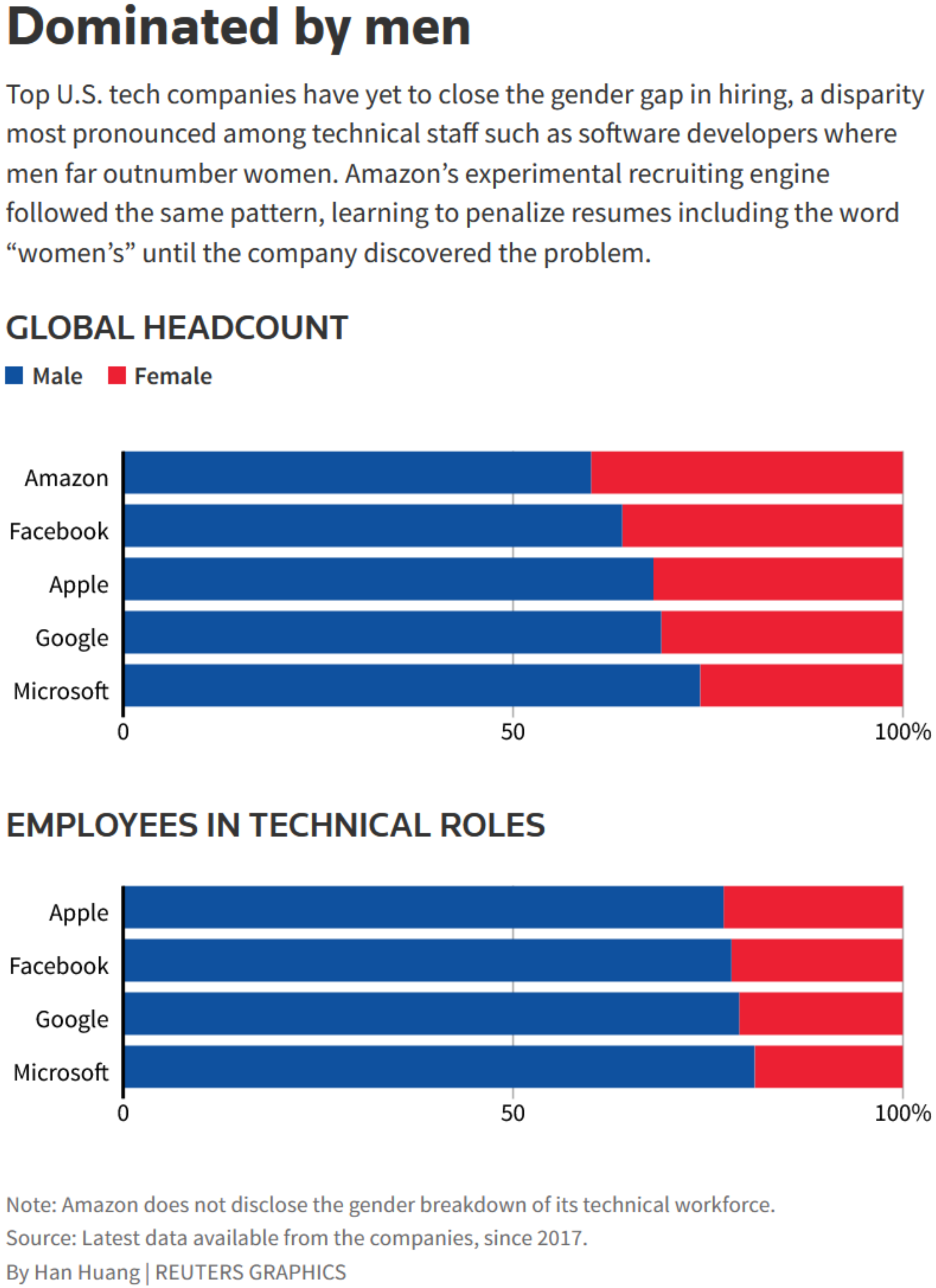
Amazon built an AI tool to hire people but had to shut it down because it was discriminating against women

Isobel Asher Hamilton Oct 10, 2018, 5:47 AM

Amazon CEO Jeff Bezos. David Ryder/Getty Images

- **Amazon tried building an artificial-intelligence tool to help with recruiting, but it showed a bias against women, Reuters reports.**
- **Engineers reportedly found the AI was unfavorable toward female candidates because it had combed through male-dominated résumés to accrue its data.**
- **Amazon reportedly abandoned the project at the beginning of 2017.**

(https://www.businessinsider.com/amazon-built-ai-to-hire-people-discriminated-against-women-2018-10)



Dominated by men

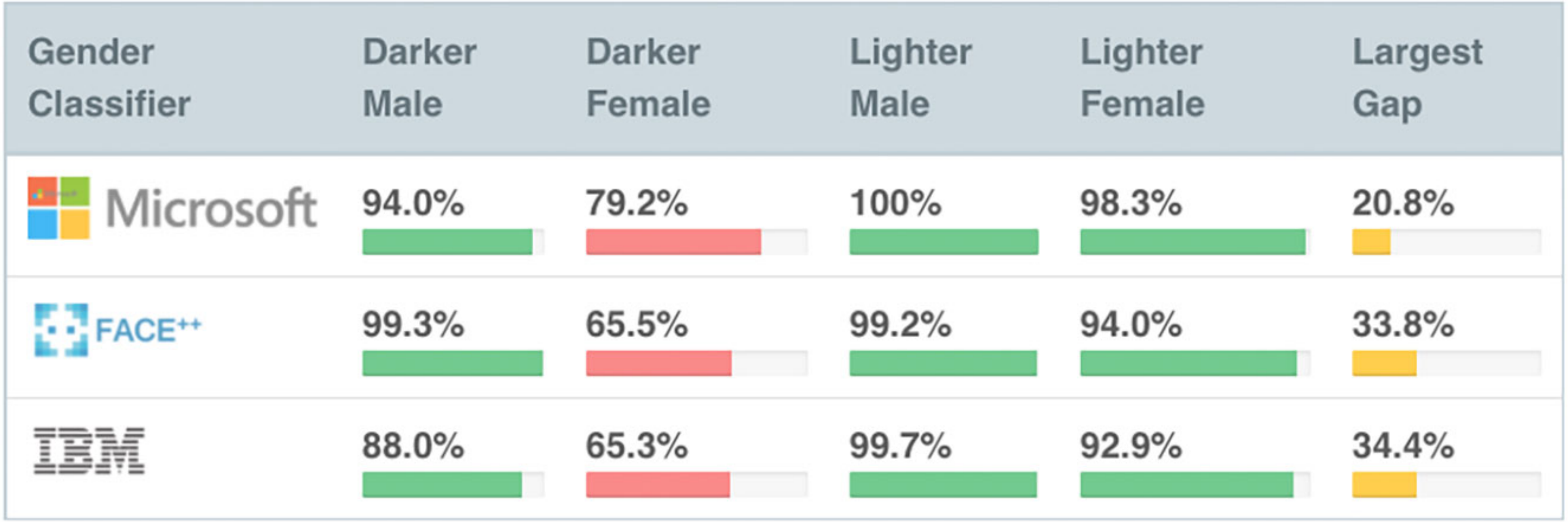
Top U.S. tech companies have yet to close the gender gap in hiring, a disparity most pronounced among technical staff such as software developers where men far outnumber women. Amazon's experimental recruiting engine followed the same pattern, learning to penalize resumes including the word "women's" until the company discovered the problem.

GLOBAL HEADCOUNT

🟦 Male 🟥 Female

EMPLOYEES IN TECHNICAL ROLES

Note: Amazon does not disclose the gender breakdown of its technical workforce.
Source: Latest data available from the companies, since 2017.
By Han Huang | REUTERS GRAPHICS



Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

SOME OTHER ISSUES… (NON-EXHAUSTIVE)

Privacy: where is our data coming from, and how does this affect the privacy of our users?

Autonomy: do our algorithms encroach on people’s autonomy, e.g. by nudging them to make the choices we want them to, or encouraging them to let the algorithm make choices for them? Is this a problem?

Social Issues: do our algorithms cause any broad social problems? (Example: content recommendation might reinforce pre-existing but problematic social categories)

LAW AND MORALITY

LAW VS MORALITY

LAW VS MORALITY

LAW VS MORALITY

- What is legal is not necessarily moral.
- What is illegal is not necessarily immoral.
    - Although in a just society, you should almost always follow the law.
    - Also, in areas of well-developed law, the law can be a good but fallible source of moral insight.
- Legal compliance should be thought of as the minimum standard for responsible conduct.
    - This is especially true for emerging technology fields, where legal standards may be poorly developed.

FAIRNESS AND DISCRIMINATION

DISPARATE TREATMENT

VS.

DISPARATE IMPACT

- Title VII of the Civil Rights Act provides a particularly robust set of anti-discrimination laws pertaining to employment.
  - Aims to protect certain “protected classes” from certain forms of discrimination.
  - Provides a useful conceptual framework for approaching discrimination in machine learning.

Protected Class: e.g. race, religion, national origin, age, gender, pregnancy status, disability status, veteran status …

Disparate Treatment: roughly, intent to discriminate, i.e. intentionally treating someone unequally because of their membership in a protected class.

Disparate Impact: policies or practices that, while formally non-discriminatory, have disproportionate and adverse effects on certain protected classes.

DISPARATE TREATMENT

VS.

DISPARATE IMPACT

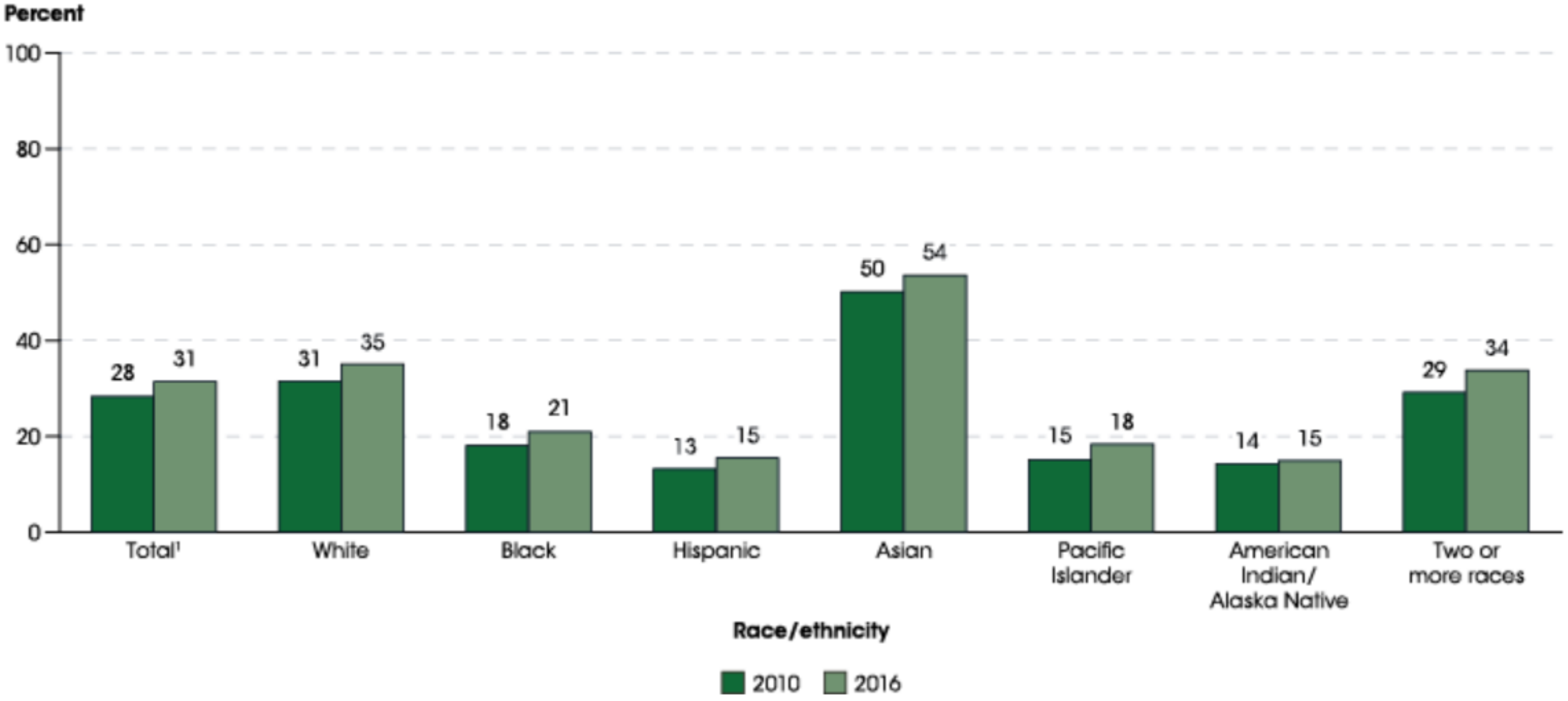
Disparate Treatment: roughly, intent to discriminate, i.e. intentionally treating someone unequally because of their membership in a protected class.

Disparate Impact: policies or practices that, while formally non-discriminatory, have disproportionate and adverse effects on certain protected classes.

An example of disparate *impact* that does not involve disparate *treatment*:

- Imagine I’m hiring somebody to work as a cashier in my store.
- I decide I will only hire somebody with a college degree.
- This does not, on its face, involve disparate treatment.
- But it likely will involve disparate impact, because having a college degree correlates strongly with race.

Figure 27.3. Percentage of adults age 25 and older who had completed a bachelor's or higher degree, by race/ethnicity: 2010 and 2016



[1] Total includes other racial/ethnic groups not separately shown as well as respondents who wrote in some other race that was not included as an option on the questionnaire and therefore could not be placed into any of the other groups.

NOTE: Race categories exclude persons of Hispanic ethnicity. Although rounded numbers are displayed, the figures are based on unrounded estimates.

SOURCE: U.S. Department of Commerce, Census Bureau, American Community Survey, 2010 and 2016. See *Digest of Education Statistics 2017*, table 104.40.

DISPARATE TREATMENT

VS.

DISPARATE IMPACT

The "Business Necessity" Defense: sometimes, as a legal matter, disparate impact can be justified as a "business necessity".

"A challenged employment practice must be "shown to be related to job performance," have a "manifest relationship to the employment in question," be "demonstrably a reasonable measure of job performance," bear some "relationship to job-performance ability," and/or "must measure the person for the job and not the person in the abstract." (Griggs v. Duke Power Co, as quoted in Barocas and Selbst 2016)

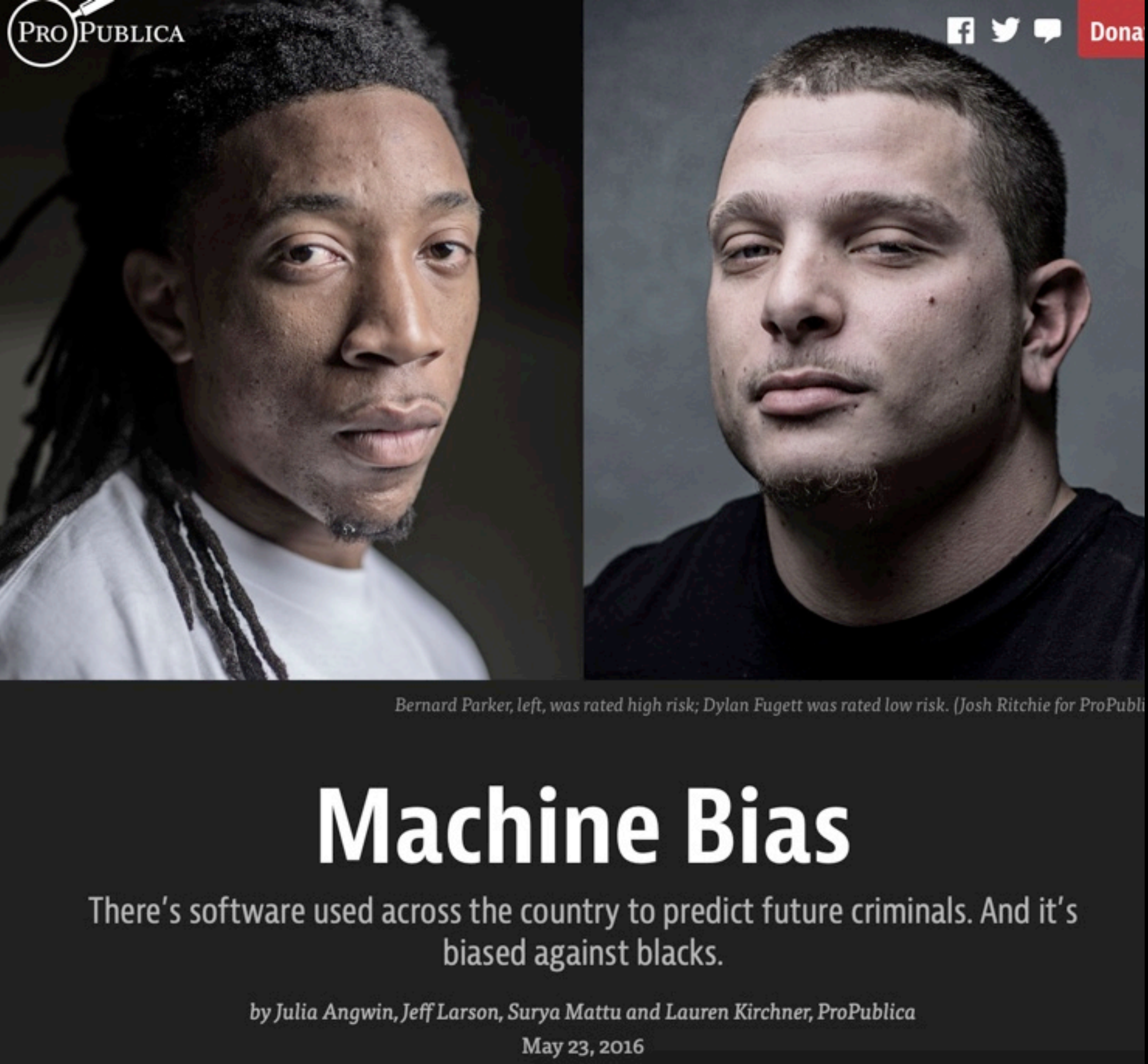
A fictitious example:

- I work for a medical clinic and I'm hiring a doctor
- I will only hire people with medical degrees.
- This is very likely to have a disparate impact.
- But it is nonetheless a permissible practice, since I am not trying to discriminate, and there are legitimate reasons for this disparate impact.

DISCRIMINATION IN MACHINE LEARNING ALGORITHMS

Machine learning algorithms can produce disparate impact even without disparate treatment.

Disparate impact can be a result of bad data practices, or even an inevitable product of data mining itself.



PRO PUBLICA

Dona

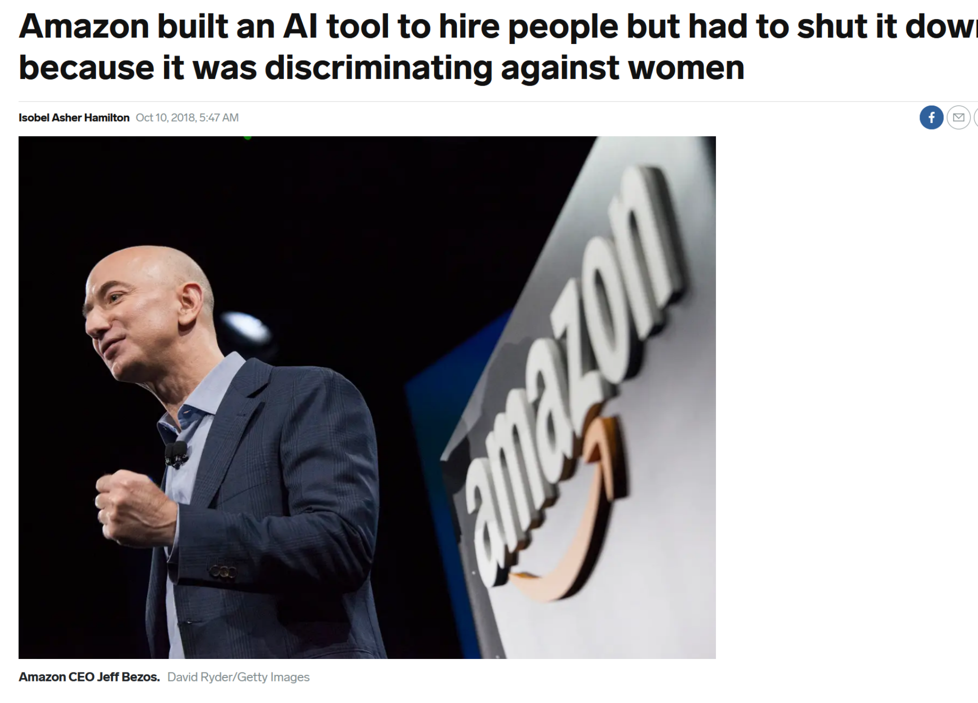
Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPubli

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

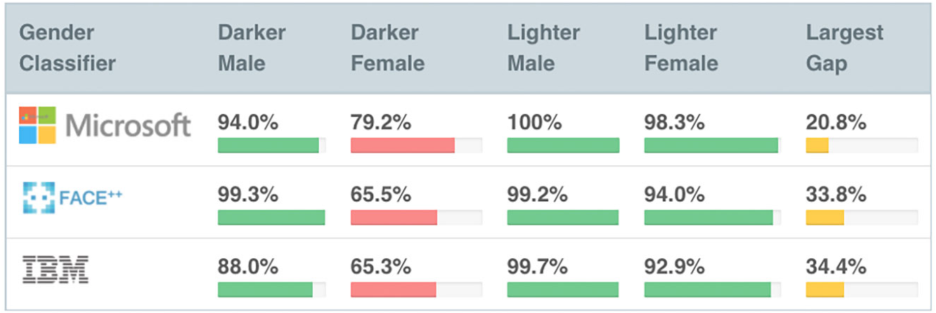
May 23, 2016



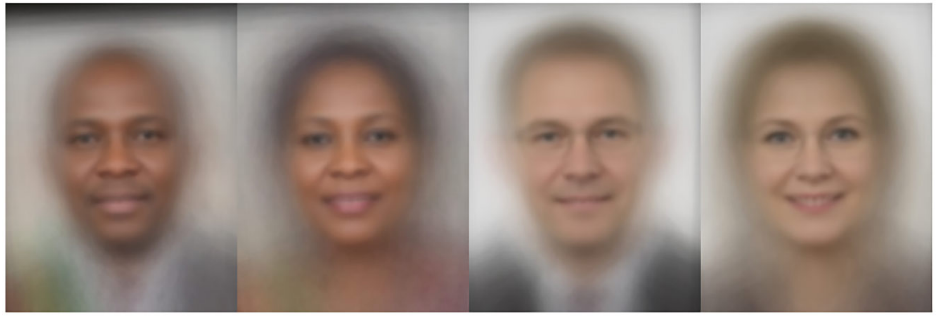
Amazon built an AI tool to hire people but had to shut it dow because it was discriminating against women

Isobel Asher Hamilton Oct 10, 2018, 5:47 AM

Amazon CEO Jeff Bezos. David Ryder/Getty Images



Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

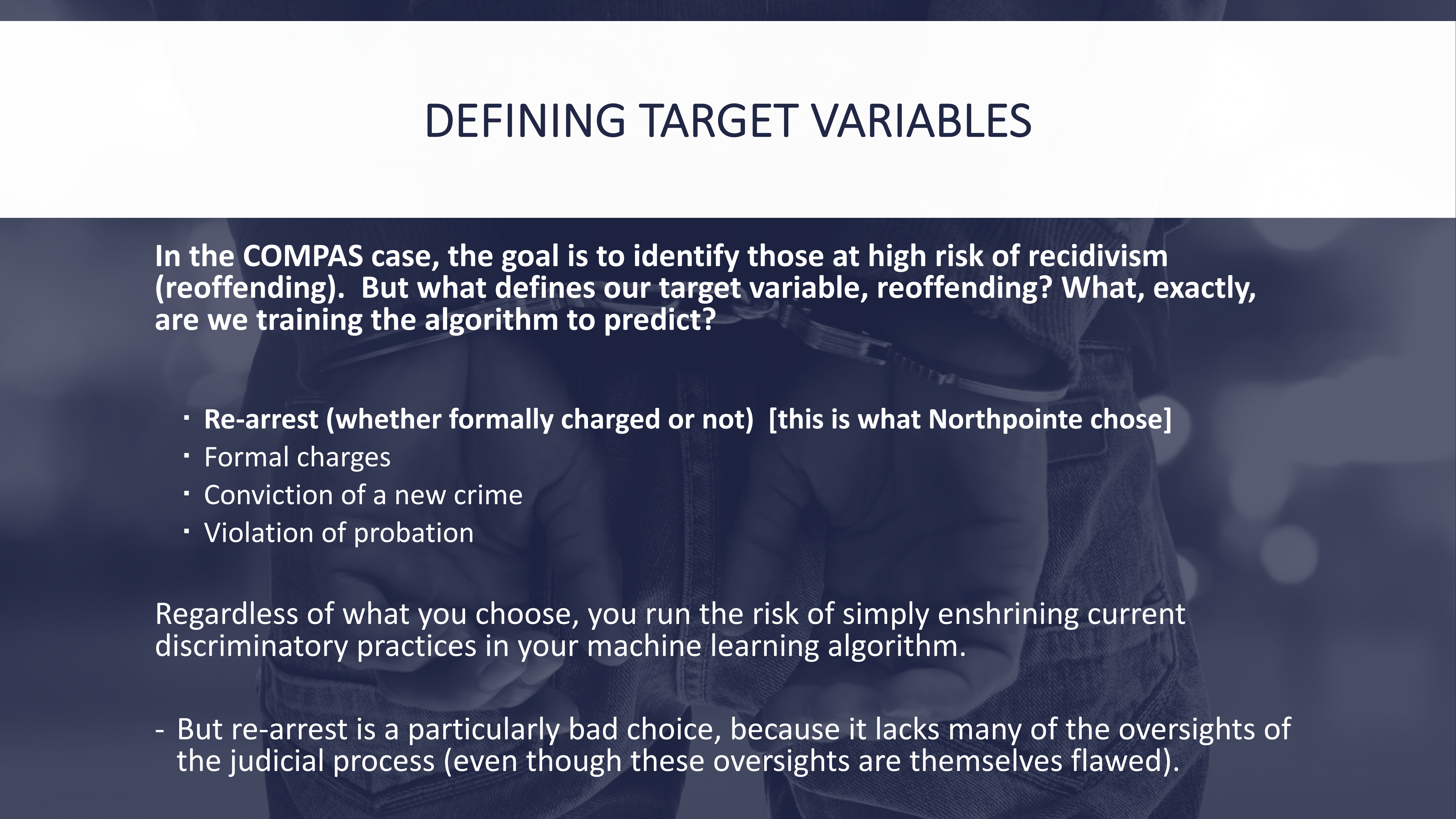


POTENTIAL SOURCES OF BIAS

- When defining target variables and in class labels
- When assembling the training data set, resulting in an unrepresentative sample
- When selecting relevant features
- Intentional bias (this is always still possible)

(This section of the presentation draws heavily from “Big Data’s Disparate Impact” (2016), by Barocas and Selbst)

DEFINING TARGET VARIABLES



In the COMPAS case, the goal is to identify those at high risk of recidivism (reoffending). But what defines our target variable, reoffending? What, exactly, are we training the algorithm to predict?

- **Re-arrest (whether formally charged or not) [this is what Northpointe chose]**
- Formal charges
- Conviction of a new crime
- Violation of probation

Regardless of what you choose, you run the risk of simply enshrining current discriminatory practices in your machine learning algorithm.

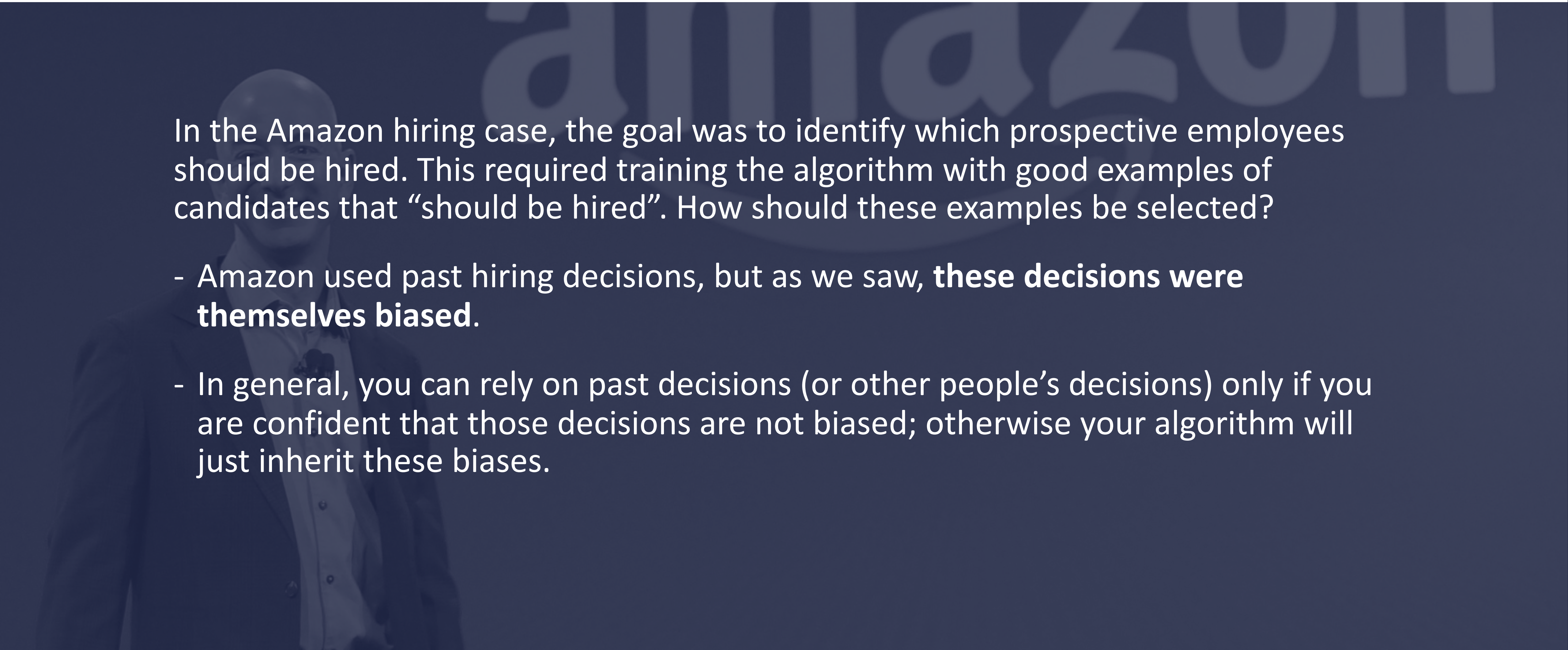
- But re-arrest is a particularly bad choice, because it lacks many of the oversights of the judicial process (even though these oversights are themselves flawed).

ASSEMBLING THE TRAINING DATA SET

The **labeling** of training examples can be corrupted by bias because it is subjective. Data that looks "objective" often fails to be so.

Data may be **inaccurate** or incomplete for certain classes of people, or it may **overrepresent** them

LABELING TRAINING EXAMPLES: ENSHRINING PAST PREJUDICE



In the Amazon hiring case, the goal was to identify which prospective employees should be hired. This required training the algorithm with good examples of candidates that “should be hired”. How should these examples be selected?

- Amazon used past hiring decisions, but as we saw, **these decisions were themselves biased**.

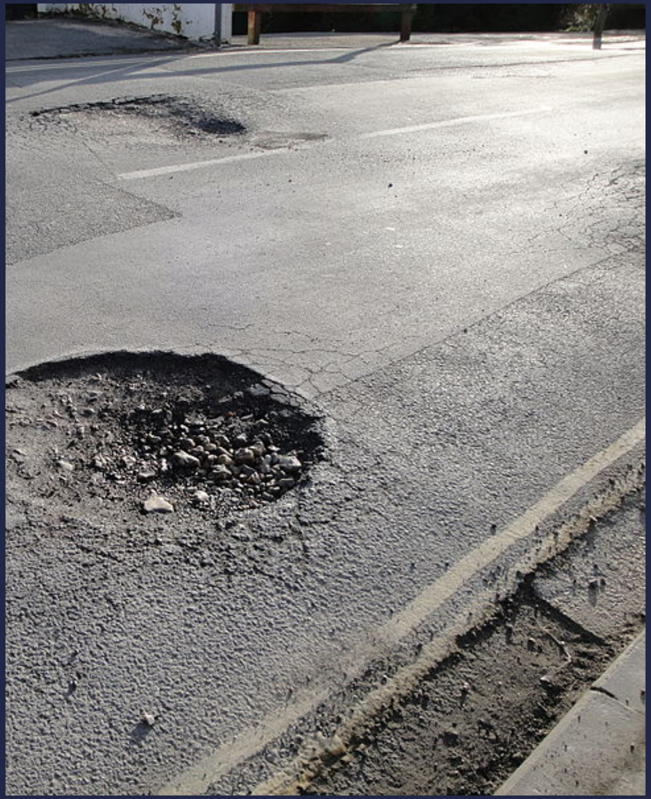
- In general, you can rely on past decisions (or other people’s decisions) only if you are confident that those decisions are not biased; otherwise your algorithm will just inherit these biases.

LABELING TRAINING EXAMPLES: REINFORCING CURRENT PREJUDICE

"A similar situation could conceivably arise on websites that recommend potential employees to employers, as LinkedIn does through its Talent Match feature. If LinkedIn determines which candidates to recommend based on the demonstrated interest of employers in certain types of candidates, Talent Match will offer recommendations that reflect whatever biases employers happen to exhibit. In particular, if LinkedIn's algorithm observes that employers disfavor certain candidates who are members of a protected class, Talent Match may decrease the rate at which it recommends these candidates to employers. The recommendation engine would learn to cater to the prejudicial preferences of employers."

(Barocas and Selbst 2016, 683)

UNDERREPRESENTATION IN DATA SETS

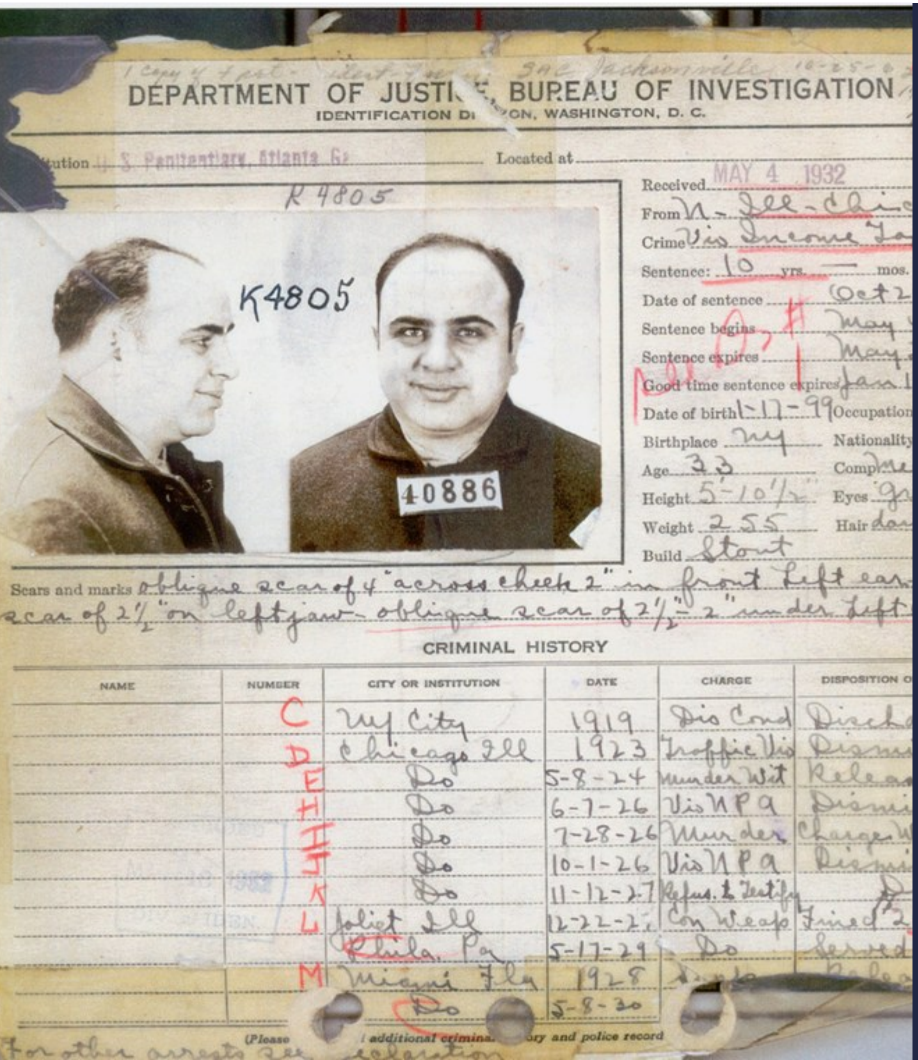


- “Street Bump” was a program that used accelerometer data from smart phones to detect potholes.
- This was used to report road problems to the city (in this case, Boston), who could then fix them.
- But affluent neighbourhoods are likely to be far better represented in this data set than poorer neighbourhoods (due to differences in smart phone ownership).

OVERREPRESENTATION IN DATA SETS

- “Overrepresentation in a dataset can also lead to disproportionately high adverse outcomes for members of protected classes. Consider an example from the workplace: managers may devote disproportionate attention to monitoring the activities of employees who belong to a protected class and consequently observe mistakes and transgressions at systematically higher rates than others, in part because these managers fail to subject others who behave similarly to the same degree of scrutiny. Not only does this provide managers with justification for their prejudicial suspicions, but it also generates evidence that overstates the relative incidence of offenses by members of these groups. Where subsequent managers who hold no such prejudicial suspicions cannot observe everyone equally, they may rely on this evidence to make predictions about where to focus their attention in the future and thus further increase the disproportionate scrutiny that they place on protected classes.” (Barocas and Selbst 2016, 687)
- **A real-world example:** blacks and whites smoke marijuana at same rates (by own admission in surveys) but blacks are 4-5 times more likely than whites to be arrested for marijuana-related offenses.

WHEN SELECTING RELEVANT FEATURES



1 Copy 4 [illegible] SAC Jacksonville 10-25-[illegible]

DEPARTMENT OF JUSTI E, BUREAU OF INVESTIGATION

IDENTIFICATION DI ON, WASHINGTON, D. C.

tution U. S. Penitentiary, Atlanta Ga Located at

K 4805

K4805

40886

Received MAY 4 1932

From N. [illegible]-Chi

Crime Vio Income Ta

Sentence: 10 yrs. — mos.

Date of sentence Oct 2

Sentence begins May

Sentence expires May

Good time sentence expires Jan 1

Date of birth 1-17-99 Occupation

Birthplace NY Nationality

Age 33 Comp Me

Height 5'-10½" Eyes Gr

Weight 255 Hair da

Build Stout

Scars and marks oblique scar of 4" across cheek 2" in front left ear scar of 2½" on left jaw - oblique scar of 2½" - 2" under left

CRIMINAL HISTORY

NAME	NUMBER	CITY OR INSTITUTION	DATE	CHARGE	DISPOSITION O
	C	NY City	1919	Dis Cond	Disch
	D	Chicago Ill	1923	Traffic Vio	Dismi
	E	Do	5-8-24	Murder Wit	Releas
	H	Do	6-7-26	Vio NPA	Dismi
	I	Do	7-28-26	Murder	Charge W
	J	Do	10-1-26	Vio NPA	Dismi
	K	Do	11-12-27	Refus. to Testify	[illegible]
	L	Joliet Ill	12-22-2	Con Weap	Fined 2
		Phila. Pa	5-17-29	Do	Served
	M	Miami Fla	1928	[illegible]	Releas
		Do	5-8-30		

(Please additional crimina ory and police record

For other arrests see [illegible]

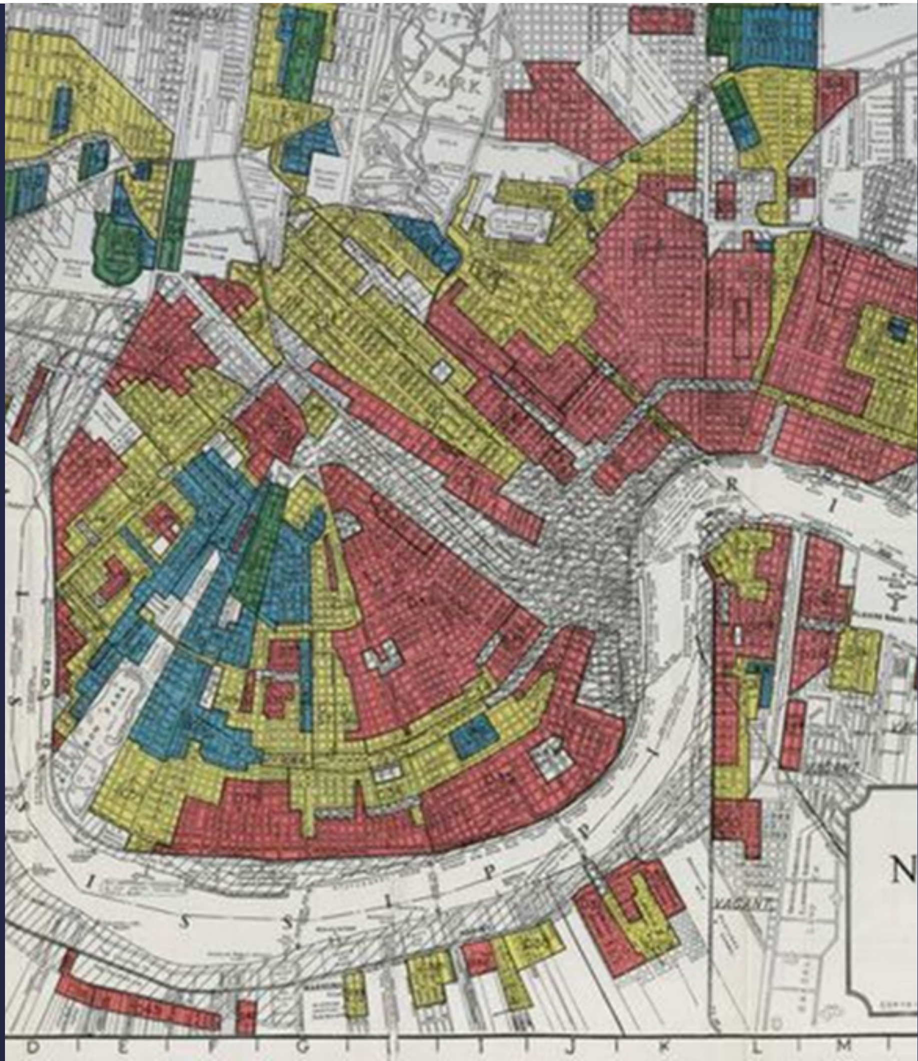
Ex in COMPAS case, which features of data should we select (ie, what data should we collect)?

- Criminal history
- Age at first arrest
- Gender
- Socio-economic status
- Current employment status
- Criminal companions/associations
- Antisocial behavior
- Family criminality

INTENTIONAL BIAS IN DATA COLLECTION

Data mining can obscure intentional discrimination

- What can be done accidentally can also be done intentionally.
  - (Somebody could train a machine learning algorithm on past training data *because they know* this data is biased and want a biased hiring algorithm).
- Data mining can enable institutions to circumvent legal barriers to unlawful discrimination by making proof very difficult to obtain



VALUE SENSITIVE DESIGN

(see: vsd.ccs.northeastern.edu for more information)

VALUES AND TECHNOLOGY

Technology is the result of human imagination

All technology involves design

All design involves choices among possible options

All choices reflects values

Therefore, all technologies reflect and affect human values

Ignoring values in the design process is irresponsible

Making ethical judgments is sometimes an unavoidable part of producing technology, and the stakes can be high!

VSD IN ACTION: SOME CORE COMPONENTS

1. Identify stakeholders.
2. Identify the values at stake for these stakeholders.
3. Discard illegitimate values.
4. Identify where value tradeoffs are necessary.
5. Prioritize important values.
6. Use this to define success.

Note: this is not a foolproof decision procedure; it is a heuristic device to help you recognize the values that are relevant to your design project, and to respond to them appropriately.

VSD IN ACTION: SOME CORE COMPONENTS

1. **Identify stakeholders.**
2. Identify the values at stake for these stakeholders.
3. Discard illegitimate values.
4. Identify where value tradeoffs are necessary.
5. Prioritize important values.
6. Use this to define success.

STAKEHOLDERS IN MACHINE LEARNING

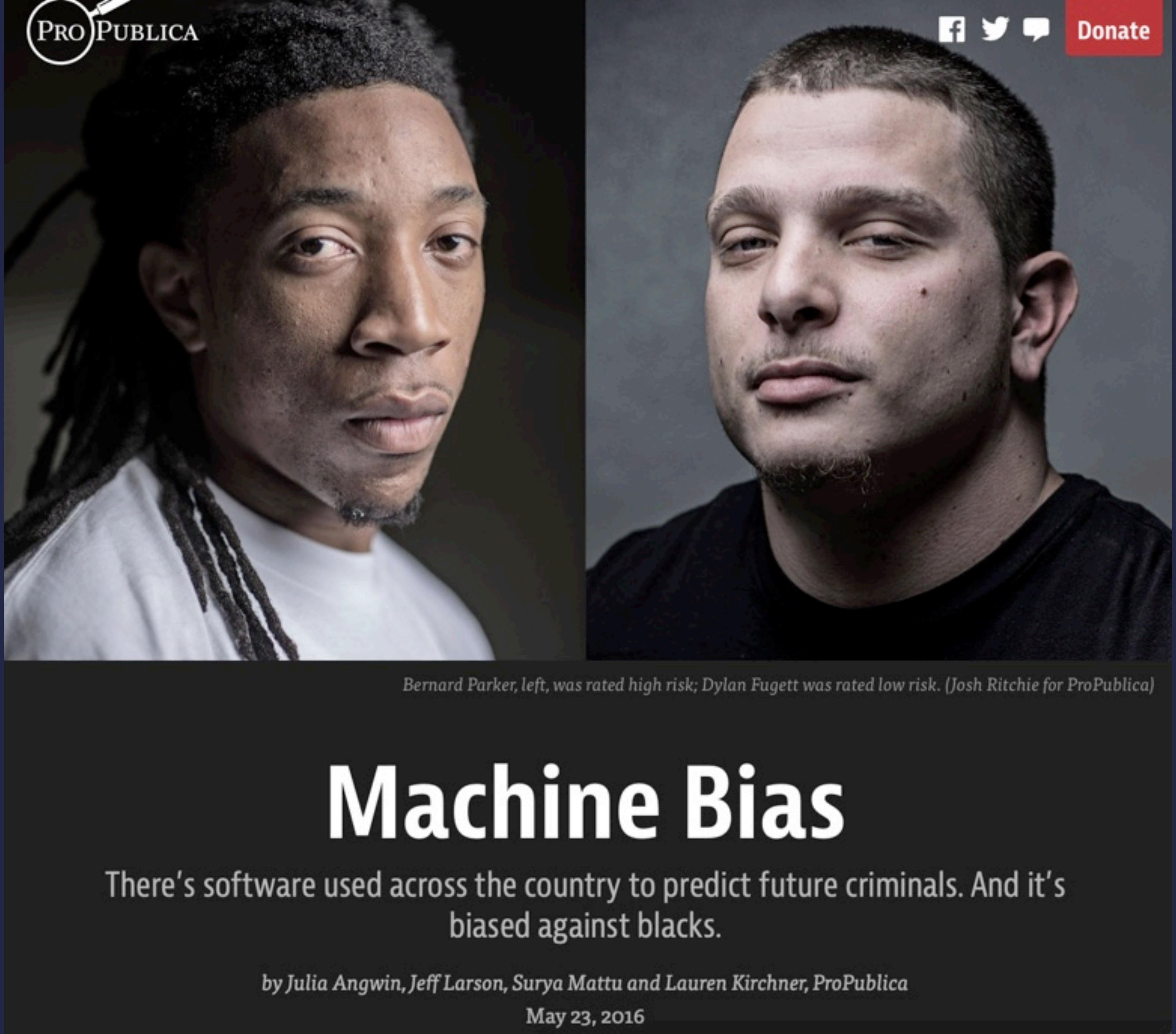
- When thinking about the values relevant to some project, it can be useful to think about whose values / interests are affected by the technology in question.

- These are the stakeholders.

  - Direct stakeholders: people who are directly impacted by the technology; typically this is users (or people it is used on), or owners.

  - Indirect stakeholders: people who do not directly interface with the technology in question, but are affected by it nonetheless.

    - This distinction is really just a heuristic; the basic point is to recognize that technology affects more than just the people who themselves use it.



Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

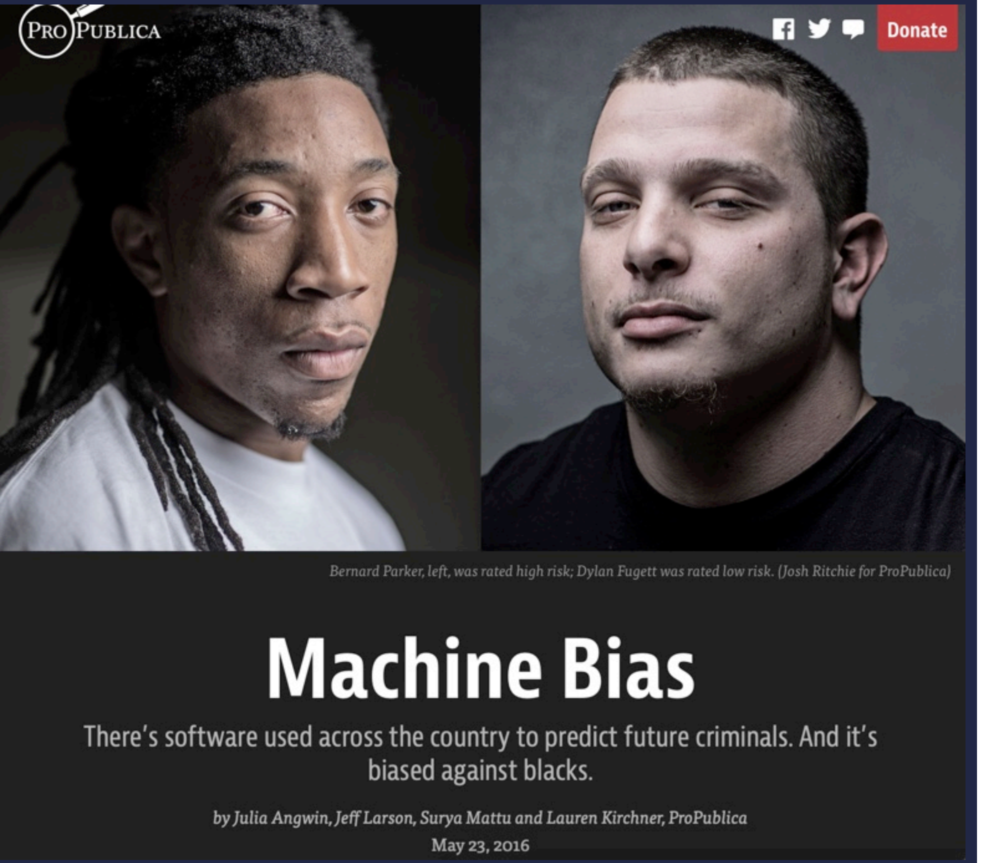
Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

STAKEHOLDERS IN COMPAS



Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

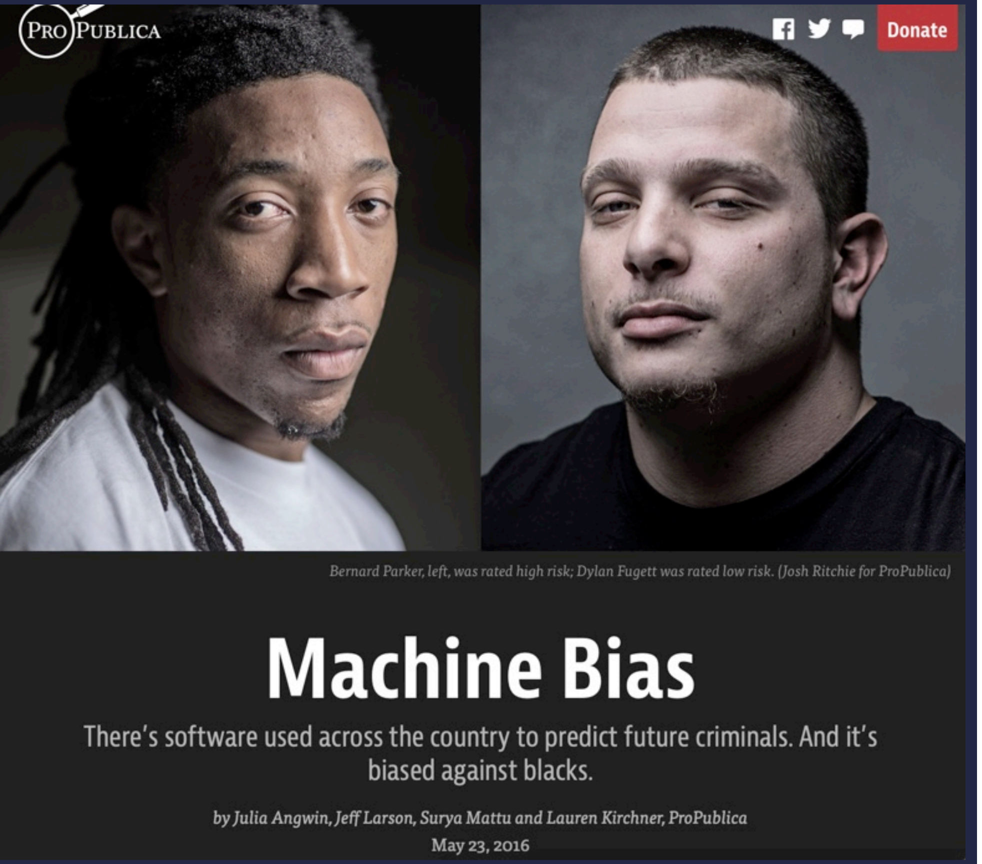
by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

- Direct Stakeholders:
    - The government / judicial system
    - People who are arrested
- Indirect Stakeholders:
    - American citizens

VSD IN ACTION: SOME CORE COMPONENTS

1. Identify stakeholders.
2. **Identify the values at stake for these stakeholders.**
3. Discard illegitimate values.
4. Identify where value tradeoffs are necessary.
5. Prioritize important values.
6. Use this to define success.

STAKEHOLDERS IN COMPAS



Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

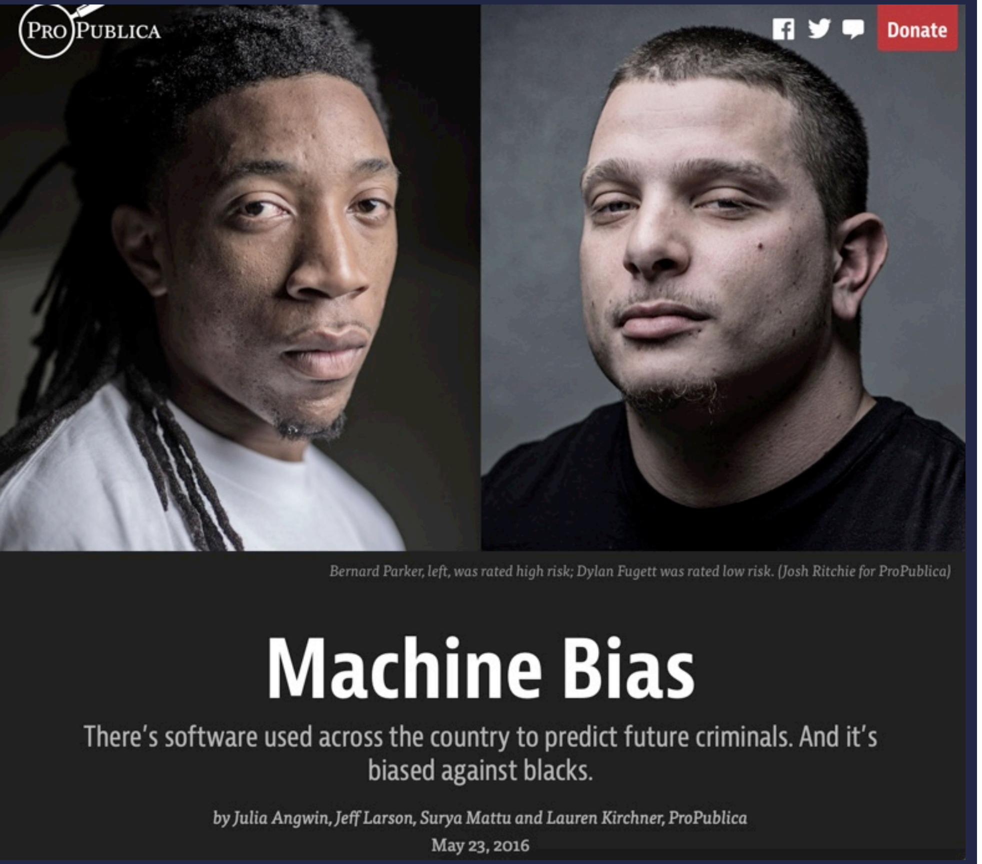
May 23, 2016

- Direct Stakeholders:
  - The government / judicial system
    - Values: efficiency; accuracy.
      - Don't want to spend unnecessary money.
      - Don't want people released on bail who are likely to recidivate.
  - People who are arrested
    - Values: freedom; right to due process; non-discrimination
      - Don't want to be detained (at all).
      - Don't want to be detained unjustly.
- Indirect Stakeholders:
  - American citizens
    - Values: similar to the government / judicial system.

VSD IN ACTION: SOME CORE COMPONENTS

1. Identify stakeholders.
2. Identify the values at stake for these stakeholders.
3. **Discard illegitimate values.**
4. Identify where value tradeoffs are necessary.
5. Prioritize important values.
6. Use this to define success.

STAKEHOLDERS IN COMPAS



Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

- Direct Stakeholders:
  - The government / judicial system
    - Values: efficiency; accuracy.
      - Don't want to spend unnecessary money.
      - Don't want people released on bail who are likely to recidivate.
  - People who are arrested
    - Values: freedom; right to due process; non-discrimination
      - ~~Don't want to be detained (at all).~~
      - Don't want to be detained unjustly.
- Indirect Stakeholders:
  - American citizens
    - Values: similar to the government / judicial system.

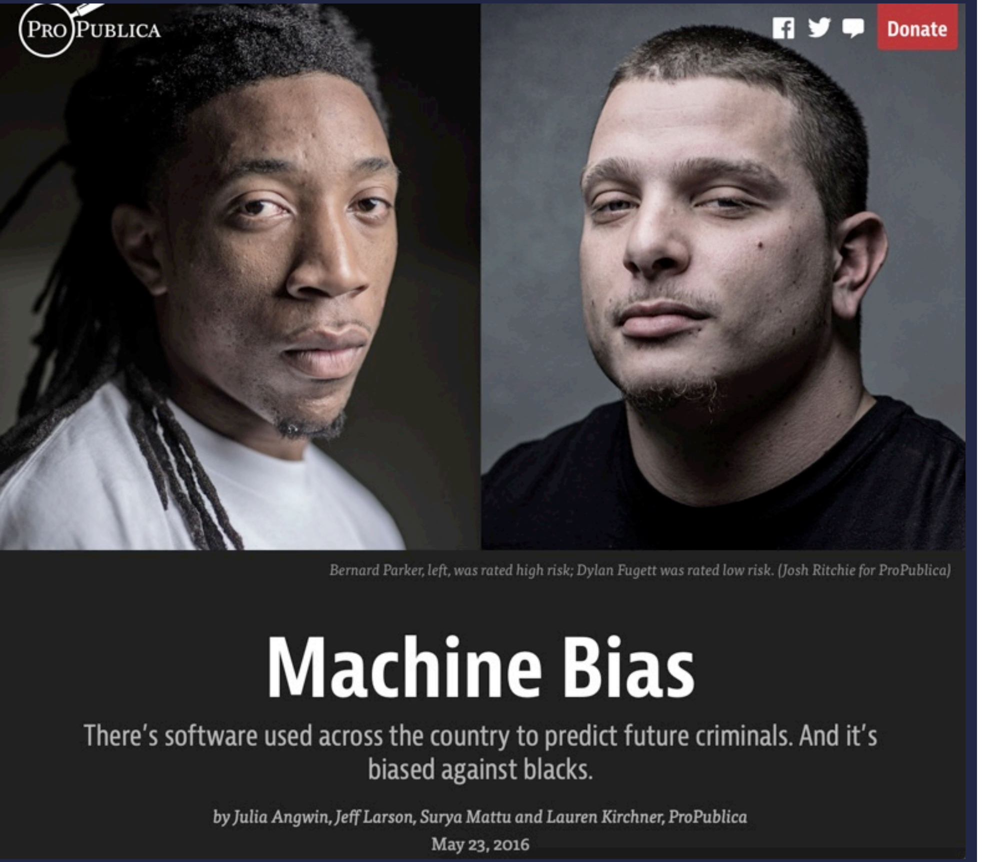
VSD IN ACTION: SOME CORE COMPONENTS

1. Identify stakeholders.
2. Identify the values at stake for these stakeholders.
3. Discard illegitimate values.
4. **Identify where value tradeoffs are necessary.**
5. Prioritize important values.
6. Use this to define success.

VALUE TRADEOFFS

- Value tradeoff: when two values are to some extent mutually incompatible, and a balance must be struck between them.
- Designing technology in a morally appropriate way means striking the (morally) right balances between all these conflicting values.
  - This is hard!
- One helpful strategy is to ask: are certain values *non-negotiable*, i.e. the kind of value that cannot be legitimately traded off against other values?

STAKEHOLDERS IN COMPAS



Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

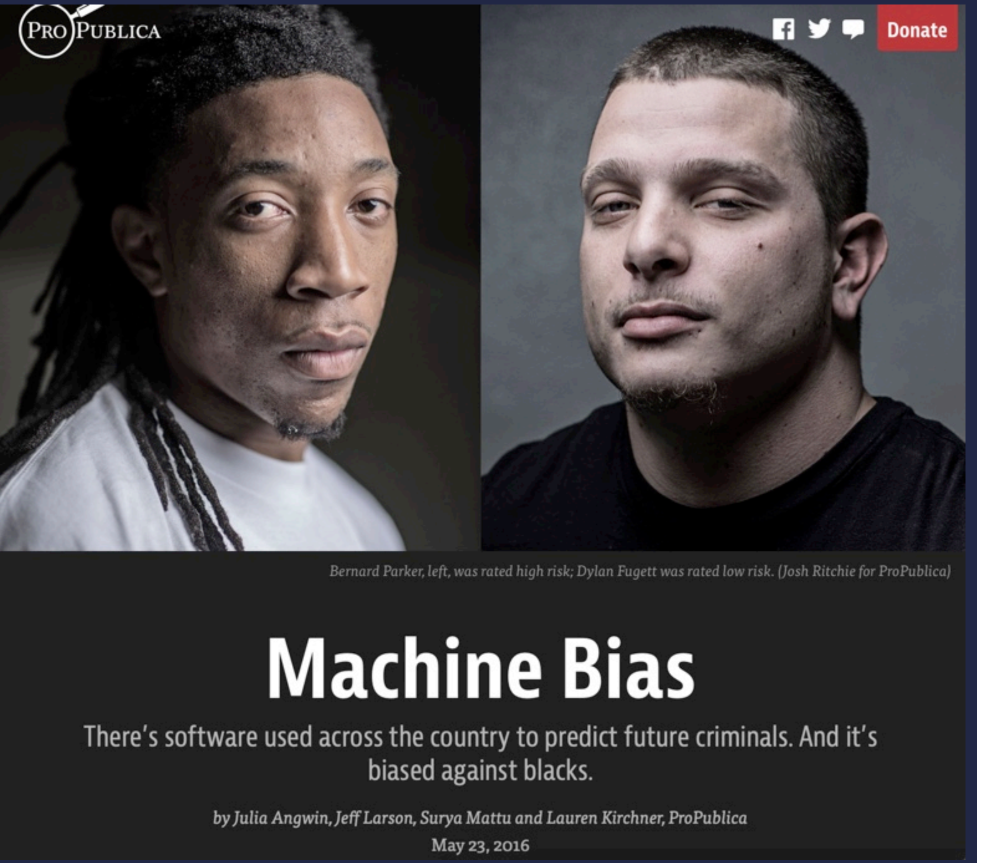
by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

- Direct Stakeholders:
  - The government / judicial system
    - Values: efficiency; accuracy.
      - **Don't want to spend unnecessary money.**
      - **Don't want people released on bail who are likely to recidivate.**
  - People who are arrested
    - Values: freedom; right to due process; non-discrimination
      - ~~Don't want to be detained (at all).~~
      - **Don't want to be detained unjustly.**

VSD IN ACTION: SOME CORE COMPONENTS

1. Identify stakeholders.
2. Identify the values at stake for these stakeholders.
3. Discard illegitimate values.
4. Identify where value tradeoffs are necessary.
5. **Prioritize important values.**
6. Use this to define success.

STAKEHOLDERS IN COMPAS



Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

- Direct Stakeholders:
    - The government / judicial system
        - Values: efficiency; accuracy.
            - **Don't want to spend unnecessary money.**
            - **Don't want people released on bail who are likely to recidivate.**
    - People who are arrested
        - Values: freedom; right to due process; non-discrimination
            - ~~Don't want to be detained (at all).~~
            - **Don't want to be detained unjustly.**

VSD IN ACTION: SOME CORE COMPONENTS

1. Identify stakeholders.
2. Identify the values at stake for these stakeholders.
3. Discard illegitimate values.
4. Identify where value tradeoffs are necessary.
5. Prioritize important values.
6. **Use this to define success.**

VALUE SENSITIVE DESIGN: DEFINING SUCCESS

For example, what exactly is the right “success” definition for COMPAS?

1. Better-than-chance predictions of who will be re-arrested, made by a system that is calibrated across racial demographics. (This is what they actually accomplished)

2. Reducing racial bias in the judicial system? (One could have this aim in designing such an algorithm)

3. Reducing biases in the judicial system more generally?

4. Reducing recidivism? (Probably not a great definition in itself … denying everybody bail will reduce recidivism in this case.)

Formulating the right “success” definition requires understanding the values that are at play in the domain in question.

VALUE SENSITIVE DESIGN: DEFINING SUCCESS

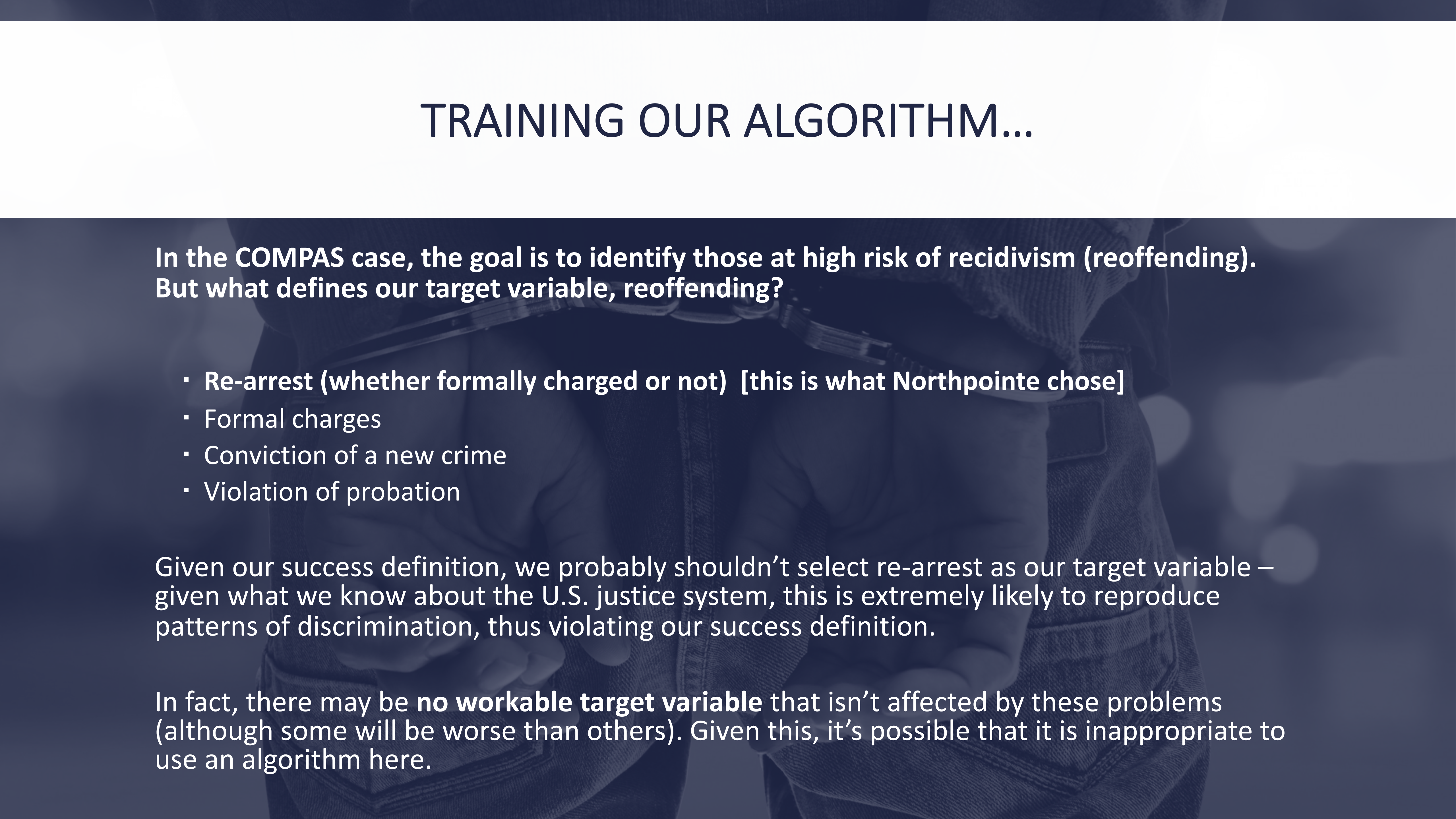
Given the values we just saw, a reasonable definition of success for COMPAS is:

- A cost-efficient algorithm
- That is (at least) better at predicting recidivism than alternate methods
- That does not detain people unjustly.

  (But this last step needs work, because what does it mean to detain somebody unjustly? Is a higher false positive rate constitutive of injustice? Is it evidence of injustice?)

  A better proposal: that does not discriminate against people.

TRAINING OUR ALGORITHM…



In the COMPAS case, the goal is to identify those at high risk of recidivism (reoffending). But what defines our target variable, reoffending?

- **Re-arrest (whether formally charged or not) [this is what Northpointe chose]**
- Formal charges
- Conviction of a new crime
- Violation of probation

Given our success definition, we probably shouldn’t select re-arrest as our target variable – given what we know about the U.S. justice system, this is extremely likely to reproduce patterns of discrimination, thus violating our success definition.

In fact, there may be **no workable target variable** that isn’t affected by these problems (although some will be worse than others). Given this, it’s possible that it is inappropriate to use an algorithm here.

AN ACTIVITY

- Choose one of these examples:
  - Personalization algorithm that selects content, on an individualized basis, in response to a search query (e.g. at Google).
  - Creditworthiness evaluation service that is used to evaluate the risk of a person defaulting on loans, e.g., credit card or mortgage debt.
  - Music recommender algorithm that selects music to recommend to users on a streaming platform (e.g. Spotify)
- Then identify:
  - 1. Stakeholders
  - 2. The legitimate values for these stakeholders.
  - 3. What a reasonable success definition would be for this algorithm.
  - 4. How this success definition might inform your data practices vis-à-vis this algorithm.